

# First Sketch of what could be the SMOS Processing scheme

Draft document to be improved

## 1. Introduction

This document aims at being a skeleton of what could be the SMOS processing philosophy. It is thus a first draft of how the SMOS ground segment could be organised. It is intended to cover all aspects from raw telemetry to final and elaborated products, e.g. not limited to the famous level 1 b<sup>1</sup>. It might turn out that the first steps are carried out at the main facilities while more elaborated products could be produced by dedicated expert centres. Tentatively these centres could focus on image reconstruction and calibration (research and improvements), soil moisture; ocean surfaces, cryosphere.

If the first part of the ground segment (the core where raw telemetry is transformed into calibrated brightness temperatures) is in principle covered by ESA in the framework of the Earth explorer Opportunity missions, the second part is not yet well ascertained. We would currently suggest that the second part is covered by structuring the European community in Thematic centres (Soil moisture, Ocean salinity, Cryosphere and eventually image reconstruction /calibration) so as to use "state of the art" algorithms to generate and disseminate elaborated products. These centres could be either co-located with the PMDC or within a national centre.

These centres would gather geographically around a dedicated processing/management and dissemination centre the activities of several science teams / laboratories. Within this structure the algorithms would be developed and validated to process satellite and ancillary data (including cal val campaigns and in situ measurements). Data could then be used for modelling or assimilation. The processing team would then mass process / reprocess the satellite data and archive / disseminate it to the end-users.

Such a scheme could be implemented on a national or international basis provided consistency is ensured.

## 2. Context

### 2.1 Generalities

Issues related to the state and the evolution of the Earth system is currently considered as a significant problem to be addressed. Global change issues, forecasting of extreme events such as major floods, improvement of weather forecasts and better management of water resources are consequently under scrutiny. For this purpose it is necessary to achieve a better understanding of the climate system, the water cycle so as to be able to monitor it at a global scale and achieve realistic projection of trends and potential future evolutions, including anthropic effects. To achieve these ambitious goals it is necessary to improve existing models and to have access to global data sets of surface variables.

Satellite remote sensing is a key element as it enables to gather repetitively and globally relevant surface variables. However, to be useful, such data must be acquired and processed

---

<sup>1</sup> Actually, as described in this document, we recommend to have ESA processing SMOS data up to level 2. The recommendation is strongly supported by the SMOS SAG.

over relatively long periods of time in a consistent manner. It is also important that data is made available to the science community at large as it fosters use and leads to innovative approaches as to their use. After the algorithms are well tested and validated they can be transferred to the users community.

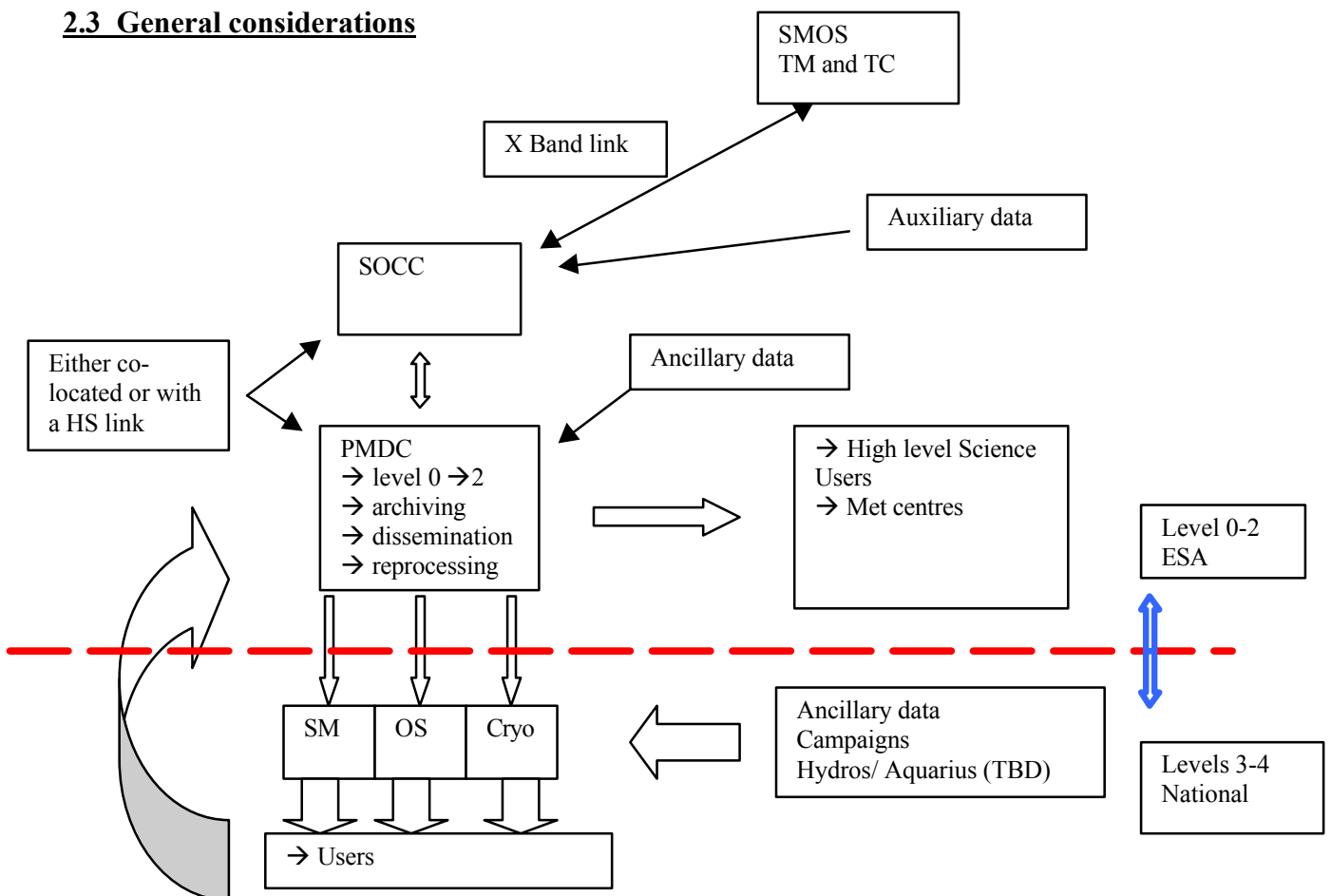
In this context, space agencies and the European Union have launched programmes aiming at the development of "observatories" of the environment based on a concerted implementation of complementary observing means (ground, A/C, and satellites) data collected by such means are to be integrated in entities able to ingest them into models for assessing and monitoring the environment. For efficiency purposes these observation means and data, their processing and analysis should be structured according to thematic considerations. It is now well admitted that, if remote sensing can provide useful information on several variables some basic ones such as soil moisture and sea surface salinity are cruelly missing.

**2.2 The SMOS case**

In all the following, we will endeavour to use classical definitions for the different levels. However, the characteristics of SMOS entail some specific features that will have to be taken into account. These will be detailed below. Not specific to SMOS but quite important is the ancillary/ auxiliary data point. To be processed efficiently such data sets will have to be available at the PMDC. It is a dimensioning factor.

In the following we will call, for the sake of clarity, **auxiliary data** what is directly linked to the mission itself whether coming directly with the telemetry (on board calibration data for instance) or coming through another path (e.g., improved satellite attitude from post processing of orbital data for a more accurate reconstruction). **Ancillary data** will correspond to the data sets necessary for processing and variable with time such as TEC, SST and LST, snow, frost , etc....

**2.3 General considerations**



As SMOS is a demonstrator programme with new techniques and measurements involved, it is expected that during the first years of life the algorithms will be improved regularly and will hence require reprocessing. This is obviously true for any programme but especially for SMOS (new instrument with surface variables never measured before). Moreover the data will be useful for long term analysis and thus the whole data set must be maintained and eventually reprocessed (very similar in some respects to what was done within the Topex Posseidon programme).

This means that the data set will have to be archived for all levels necessary for total reprocessing and that the processing speed must be adequate to avoid any bottle neck, i.e., a scene must be processed in about (rule of the thumb) a fifth of the time taken by the satellite to acquire it<sup>2</sup>. This has to be pondered by the useful operation time of the processing centres (e.g., 24/7 or only 8 hrs per working day). This is a strong requirement to be considered in the definition phases for assessing the different processing centres.

As to the delay between satellite acquisition and processed data delivery, during the initial stages, it is not assumed that near real time will be necessary and that a delay of 3 days could be acceptable<sup>3</sup>. However, with the growing interest of Met centres, and based on the ERS WSC experience, the possibility of near real time dissemination of level 1 and eventually 2 data might be expected for year two or three after launch. This is due to the fact that direct assimilation of Tbs in Met centres is very likely and we think that we should work with this assumption.

In the next sections we will give a first outline of what is expected from the different processing centres. Together with coarse flow charts and level definitions. Bearing in mind that the processing of SMOS data must be considered with three different perspectives:

- Produce Archive and disseminate SMOS data up to level 2 with reprocessing capability
- Provide an ability to send near real time data to meteorological offices
- Produce, archive, and disseminate products up to level 3 (and eventually higher) with reprocessing capability

We will not address here the communication links between the centres and only suggest possible ways of disseminating the data.

### **3 Main Processing centre**

This centre (called in SEPA PMDC) receives the telemetry from the SOCC and produces the level 2 (1b in proposal). It is expected to locate this centre in Villafranca (Spain)

The main centre will receive the telemetry, archive it and produce through the so called image reconstruction algorithm calibrated brightness temperature maps and eventually basic geophysical products.

It is assumed that this centre will process data that will be as described below.

#### **3.1 Processing steps and corresponding (tentative) data level name**

##### **3.1.1 Raw data**

This is as downlinked from the satellite, and re-transmitted through the communication network to the Payload Data Handling and Processing Centre. This might, depending on

---

<sup>2</sup> This is my estimate. It will have to be confirmed by more knowledgeable people

<sup>3</sup> The rationale here is that global coverage is obtained at the equator every 3 days and if the processing centre closes on Week ends ....

number and location of ground stations, comprise up to 6 (TBC) consecutive orbits of data, on an “as acquired” basis.[YHK1]

Raw data should probably be separated between housekeeping, calibration (auxiliary) and actual data.

The granularity is the CCSDS "packets"

This data set is only archived until the next level is produced and checked.

Special care will have to be taken in case data is acquired by different centres when merging the orbits.

### 3.1.2 Level 0 data

This is raw data but formatted into orbits (from South pole to South pole basis, assuming that calibration will be performed while over-flying the South pole), complemented by product header (orbit number, start geolocation, end geolocation, start time, end time), and all auxiliary files as needed (i.e., satellite location, attitude, instrument calibration, health). As well as indication on operation mode: data take, dual pol /full pol, calibration mode, eventually moon calibration and drifting phase in or out of moon calibration)

Note: as PROTEUS can provide time, orbit state vector, and attitude state vector to the CCSDS science packets, there might be no need for getting any extra auxiliary files (at this product level) except maybe update orbital and attitude data.

Granularity: CCSDS packets

This data set should be archived for a limited time (1 year? TBC)

### 3.1.3 level 1 a

This is the last step to produce reversible basic products:

It is on a per orbit basis (separate ascending and descending?).

Data are converted to engineering units (e.g. digital counts to temperatures, voltages, and currents). However, as most of the data are digital correlation data not processed in this step, [YHK2]the level 0 → 1 a processing might be rather trivial.

Geolocation information (satellite level) is generated and appended to each packet.

Processing from correlation products to visibilities at antenna level (noted here TBX and TBY<sup>4</sup> for the two polarisations) using calibration information.

Auxiliary information added

Necessary calibration data (in particular antenna pattern)

Granularity: packets

Archiving: This data set is the main one up to now. It should be archived for the mission lifetime at least and should contain all what is needed to go further in the processing. It is the last reversible processing (I think). The possible apodisation functions (probably fixed after initial fine tuning) would be provided together with each data product (an "Extractor Programme" which applies the (user selected) apodisation to the data "on the fly". (This has been successfully applied to other projects; applicability to SMOS needs to be discussed).

Main uses → reprocessing, test of reconstruction algorithms

### 3.1.4 level 1b Reconstruction step at instrument level

<sup>4</sup> This notation is temporary and should be made compatible with the standard SMOS reference frames notations

This is done on a per orbit basis eventually separating ascending and descending and polarisation.

Inputs:

Level 1 a  
land /sea/lakes mask

It is the step of image reconstruction for the brightness temperature maps of X and Y polarisation, at the original spatial resolution of the frames, i.e., pixel sizes and elongations as well as viewing angles vary across the swath/image.

Ideally it should be done with two or more apodisation windows, one corresponding to land and the other to oceans (TBD)

There are two options for reconstructing TX and TY angular maps in the antenna frame (director cosines):

Option 1: reconstruct TX and TY on a uniform grid

Option 2: reconstruct TX and TY in a non-uniform grid, chosen so as to correspond to a uniform geographical grid. (size better than 15 km).

In both cases: compute radiometric sensitivity. Then:

Option 1: incidence angles, pixel characteristics taken from look up tables

Compute pixel centre co-ordinates, relevant angles

Option 2 (preferred) compute and store director cosines, relevant angles, and pixel characteristics (pixel centre co-ordinates = nodes of geogr. grid taken from look up tables). It will be necessary to locate TX and TY (or 4 stokes) to take into account the displacement due to switching (1.2 sec)

Granularity: pixels in antenna reference system

Reversibility **to be checked**

Auxiliary data: best possible orbit and attitude

Ancillary data: sun and sky

Archive short term

### **3.1.5 level 1c TB on geographical grid**

Should be on an orbit basis but separated into ascending and descending as well as polarisation. Ideally the Faraday rotation correction should be done at this level.

Input:

Level 1b

Tec map

Possibly data files for next steps (TBD) i.e., atmospheric characteristics, soil maps, WS, SST, and LST.... etc...

Together with info on sun and sky

From TX and TY compute TH and TV maps

Compute covariance matrices for TH and TV

If Option 1 chosen: interpolate to geographic grid TH and TV values, pixel characteristics, incidence angles

If Option 2 : interpolation already done.

Granularity Polarised pixels

Archive total for at least 10 years

### 3.1.6 Summary

In the proposal philosophy the PMDC was to be limited to level 1-b (1-c in this text's notations) Then, the out puts of the PMDC would be level 1-a for those working on calibration and reconstruction algorithms and level 1-c for the dedicated centres and potential users such as met offices (note that even though data is at so called ground level, it is only in terms of projection, not with atmospheric corrections, though the necessary data could be included. Ideally ionospheric correction should either be performed or angle rotation indicated.

Note also that for each pixel will be a set of Tb values (angular).

#### **Yaw steering is assumed**

In this context, the product is thus TB on a lat lon grid with polarisation information at ground level (not instrument level) Calibrated, and for a set of angles.

The data format could either be on a frame basis (full available FOV, not recommended) or on a swath basis. In this case for each pixel we would have a series of Tbs a function of angle for each polarisation.

Ascending and descending orbits would be separated

The basic resolution at ground level could be 15 km all over the globe (preferred) or a quarter of a degree (about 25 km) or tailored to the temporal integration

So, organised by ascending and descending orbits, and on a lat lon grid for each imaged point (pixel and angle), the following information is contained:

Number N of measurements and for each

- Pixel value (and pol if not separated) with either faraday rotation or all elements to do it later
- radiometric sensitivity
- centre co-ordinates, lat/long
- pixel orientation: angle of long axis w.r.t. north/south direction
- pixel distance to sub track
- elongation ratio, equivalent circle diameter,
- time of acquisition
- line-of-sight angle (with the sat.)
- illumination angle (az, el) at ground level
- sun glint flag
- land/sea/mixed pixel flag (percentages)

Data gaps flagged

Flag to state whether we are in narrow or full swath?

Operation mode

Calibration info

And possibly TEC, SST; WS, LST, atmosphere characteristics

And an ancillary file with

Land use, soil maps

NB the projection grid would be lat/lon (plate caree) with either nothing (no data) or, along the considered orbit, the above mentioned

Dimensioning considerations

Ideally the centre should work continuously with a maximum processing time of less than a fifth of the actual acquisition time (to allow for reprocessing the data set when the reconstruction algorithm is improved)

The centre should be able to post calibration whenever vicarious calibration data is available (hence probably regular reprocessing).

Data dissemination.

The level 1a will only be made available to a handful of centres working on the reconstruction /calibration algorithms and locally when reprocessing is required

The Tb data set will be made available for the centres either using the data in near real time (met offices, ECMWF)

**3.2 Level 2**

We strongly suggest that level 2 is also covered within the PMDC. The rationale is as follows:

- the mission should deliver SM and OS products,
- It does not make much sense to separate the processing for land and oceans before level 2 but from level 3 on they are quite different,
- Coastal areas need to be processed over land and ocean,
- And global external /vicarious calibration will be performed partly over oceans and applied everywhere.

So we will detail below the specificities for the main targets knowing that they will be processed as one "package".

Inputs

Level 1C

Flags for big clouds/precipitations?

Flag for ionospheric storms

Calibration data

Land →      Level 2 of previous passes  
                  LST maps (Other satellite or met centres output<sup>5</sup>)  
                  Land use map including forested areas  
                  Water bodies  
                  Snow/frost? (met outputs)  
                  Topography  
                  Soils map (texture)

Ocean →      SST (sat or met outputs)  
                  WS (sat or met outputs)  
                  Faraday (is it really necessary?)  
                  Atmosphere (weather)  
                  Data from in situ measurements (network)

---

<sup>5</sup> these data sets corresponding to fast time varying quantities should be available for the time of overpass with a possible need for temporal interpolation as they are often available at UTC time. Seasonal quantities should be updated on a monthly basis. Their granularity would probably be on a pixel basis and global estimates.

**A ) Land processing**

The goal here is to produce soil moisture maps. The basis is still the per orbit (90 S -90 N and 90N - 90 S) on a lat lon grid. For the two step approach it is necessary to have data from previous passes.

The inversion is done with three parameters; error estimates are stored with results on all cells having sufficient data.

*NB: eventually calibration will be updated whenever possible. Is there a need for some sort of filtering of the data as a function of spatial resolution?*

**B) Ocean processing**

I am not very sure of what could be done here. Obviously remaining with pixel values does not make much sense as we want to do spatio temporal retrievals

We need more inputs from the ocean community.

The idea is to use OS retrievals merged with in situ data to improve calibration

Processing

Sum TBh and Tbv (faraday)?

Create a file with SST, WS in situ data and atmospheric corrections?

Do further averaging or keep all angular values knowing that it will require contiguous orbits to merge data?

Perform inversion along dwell lines with error bars<sup>[YHK3]</sup>? One should be aware that "high resolution"<sup>[YHK4]</sup> data might be useful near the coast/estuaries....

Compute mean pixel size

**C) Coastal area processing**

For pixels located along or near cost lines a mixed pixel approach will have to be taken along the general lines and flagged accordingly:

- lots of land very little sea → land processing with "forced" sea values
- lots of sea very little land → sea processing with assumed land values or neighbouring pixel values
- In between → TBD!

Out put

2 maps every day (asc and desc)?

For each pixel

TBh and TBv at 48 and 30°?

Land /sea mask

Ancillary info used

Error estimates

Mean pixel size and elongation

Ocean Sea/coast flag

Land → Sm value

Tau value (estimated or used)

Ts value

Flags as above plus eventually one if strong difference with either expected value or previous measurements exist (implying storms, snow fall, flooding, freezing, etc.).



We can either only keep pixels which satisfy the resolution criteria or keep them all and flag accordingly.

Ocean → SSS value

Granularity pixel (lat lon grid)

Archive 3 months (?)

The EASE-GRID format and procedure could be used.

## 4. Expert (dedicated<sup>6</sup>, ...) Centres

These centres are thematically oriented. They receive the level 2 data and ancillary information to produce higher-level data products (3 to 4). They are mixing research expertise and algorithm development experts couple with mass data processing and dissemination. They should also be able to reprocess data as better calibration possibilities are made available or when improve algorithms are validated. Two such centres seem absolutely necessary (SM and OS) as the processing involved differ. A third one (cryosphere) could be considered separately or included in the land centre (as even sea ice would probably be processed similarly to land data).

We will only detail the OS and SM centres below.

### 4.1 Level 3

This corresponds to resampled and temporally accumulated data. In some cases; due to the time required to do the accumulation, it will be necessary to perform some sort of assimilation (?)<sup>[YHK5]</sup>

#### **Add parag on reconstruction**

##### **4.1.1 Land global SM Maps**

###### Inputs

3 days of level 2 (possibility to produce every day with updates on corresponding orbits?)  
useful for high latitudes mainly

Two global maps are produced (ascending and descending) on a regular grid (0.15x0.15 deg)  
At higher latitudes this does not necessarily make sense but there are more values to put in more "bins"...but the best option is probably to "filter" or average and have a coarser longitudinal resolution.

###### Output

Data very similar to that of level 2 with the addition of a flag when rain has occurred.

##### **4.1.2 Ocean**

###### Inputs

Level 2 data (day and night separated)

---

<sup>6</sup> Exact name TBD

In situ data if not already available

### Processing

Filter "bad" pixels (sun glint, islands?, storms, ....)

Build averages on GODAE elements (200 x 200 km 10 days or whatever is finally chosen)

Estimate errors

Assess sudden transients.[YHK6]

Flag "anomalies" in the temporal sense

Keep relatively high spatial resolution near coasts/estuaries

Perform calibration with in situ data

### Outputs

High resolution coarse (1 psu?) SSS maps near coast and estuaries → users and land centre

Update calibration → users and processing centres

Global maps of SSS with GODAE specs (5 day sliding maps)

### NOTES:

These are the two basic products. It is quite sure that the level 3 will be asked for other outputs such as daily temporal evolution over an area (land) different spatio temporal averaging over oceans etc... It is expected to advance on these points during future SAG meetings and SMOS workshops.

## **4.2 Level 4**

There are a number of level 4 products which one might think of. Some are fairly feasible and necessary and should thus be produced in the expert centres; some others are more research products and could be investigated outside the centres. We will give below a short overview of such products

- Assimilation of SM fields to estimate root zone soil moisture for applications in meteorology and water resources management
- Merge with optical data to use in SVATS for carbon cycle studies and flux assessment (latent and sensible) , use in vegetation growth models etc
- Statistical analysis of SM temporal evolution for climate studies, for assessing precipitations, detecting anomalies (floods, droughts)
- Produce regional maps and dis-aggregate over basins
- Assimilate SSS in Ocean Circulation Models (Mercator like)
- Use SSS for monitoring ENSO and NAO
- ....

## **5. The calibration Validation Issue**

The calibration Validation of SMOS data should be covered in the ground segment. We reckon that Validation is more of level 4 type and will left aside for the time being. For calibration we will exclude here on board calibration and concentrate on vicarious/external calibration. Currently the plans are to use several sources of information to carry out calibration. These include

- Sky measurements (eventually the moon). They will be performed regularly and could easily be implemented at level 1c for a general correction of the instrument calibration
- The "Ruf" method which could be done routinely as well and with a relatively high frequency. They could be done (TBC) at level 2
- The use of in situ measurements over the ocean which could also be done routinely and applied at level 2
- The use of data from other missions (Aquarius, Hydros) which are still TBD pending more info on these mission and whether they fly in the same timeframe.

In all cases the idea would be to improve calibration at level 1C (general correction of the calibration) or at level 2 (specific methods often statistical and latitude / target dependent) but applied everywhere. History of these calibration would be kept and eventually applied during later reprocessings as on the fly calibration with external data sets might pose some issues with the ground segment definition and applicability.

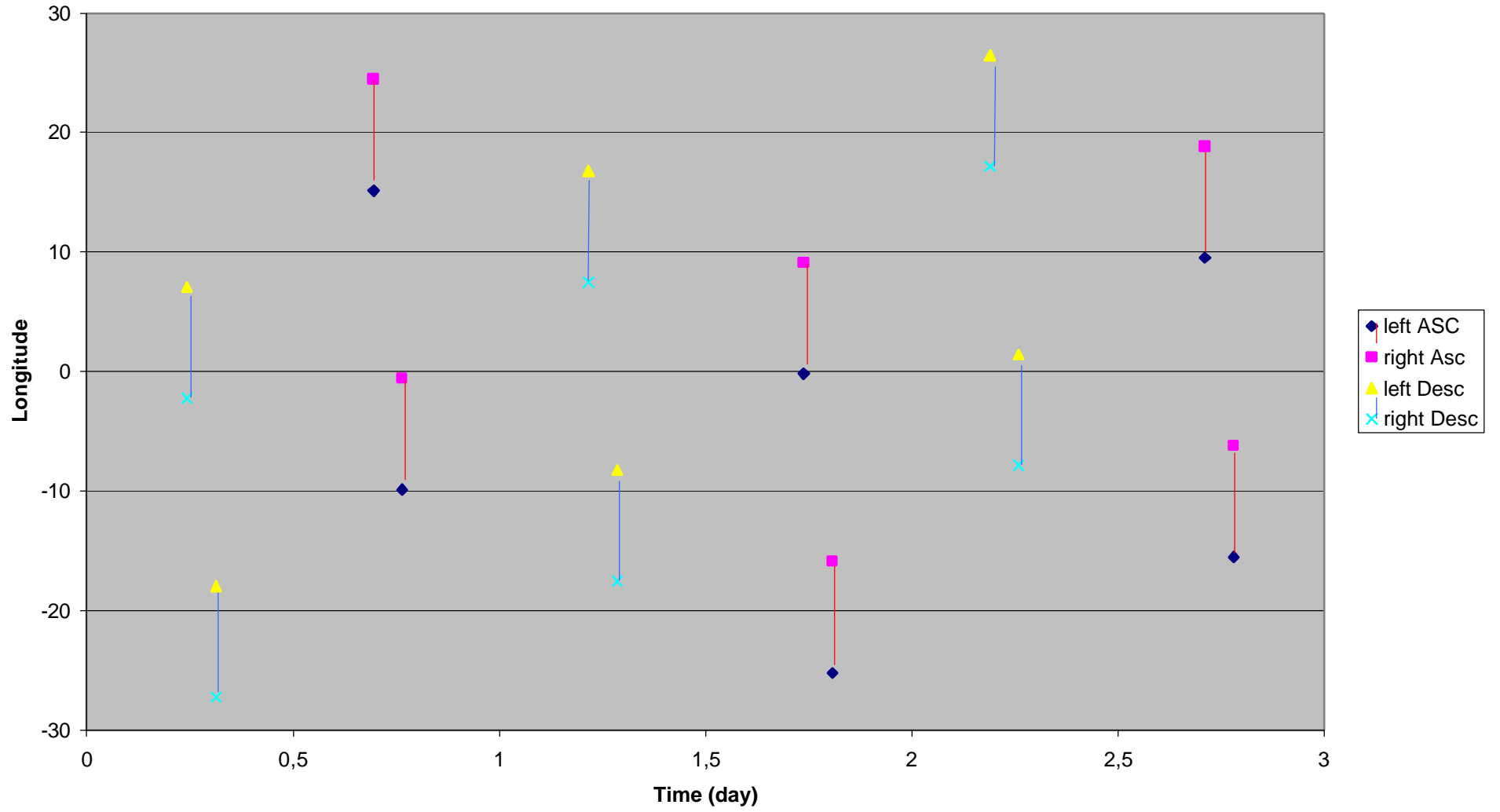
### **Annex: SMOS coverage**

So as to get a better picture of the SMOS coverage, you will find in this annex the swaths obtained at different latitudes (0; 10, 20, ..., 80°N) with both ascending and descending passes.

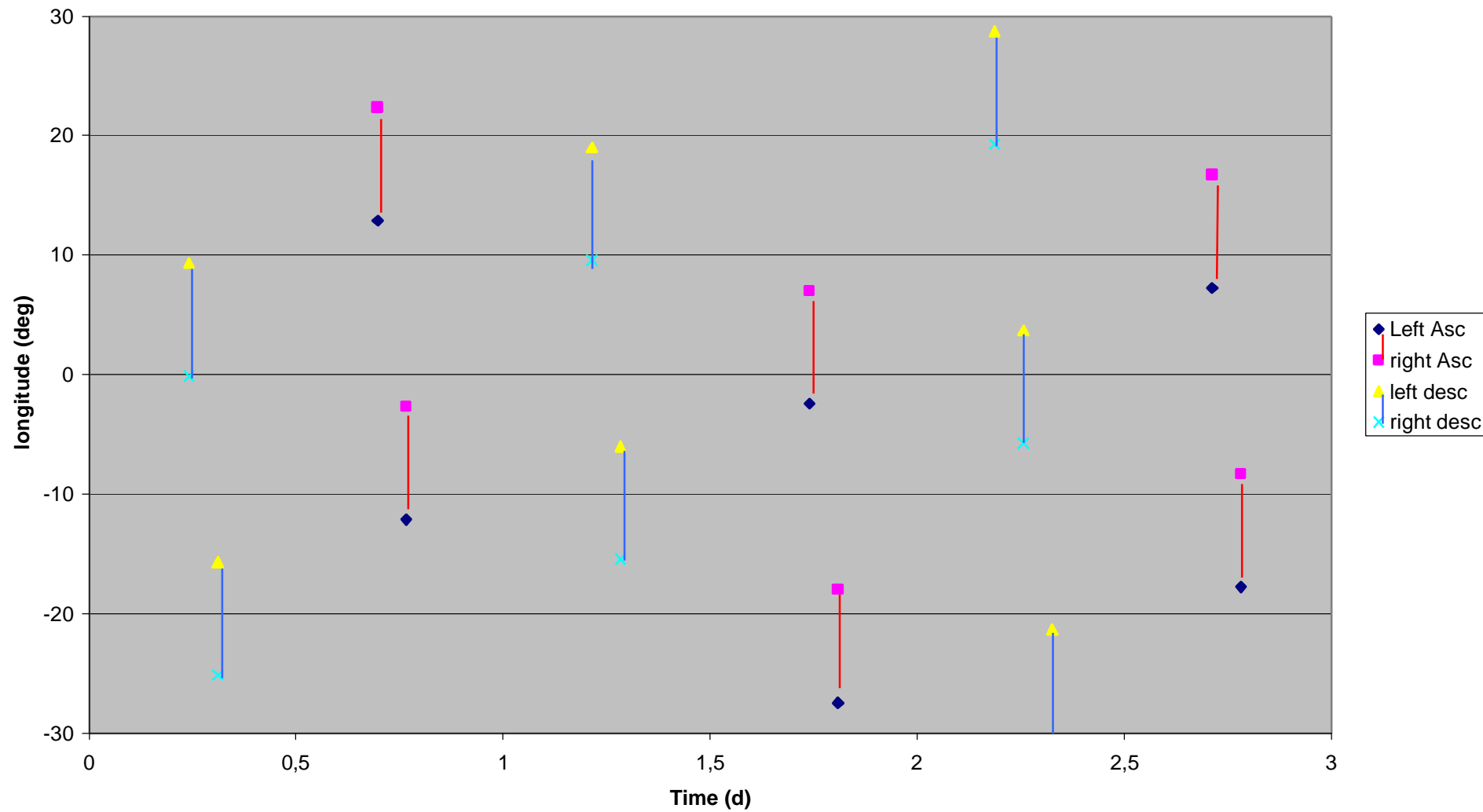
The plots show for each latitude and as a function of time (X axis) the swath (blue and red lines) expressed as longitudes. This is computed for the baseline SMOS characteristics.

I have made this with coarse tools and it is only indicative of course, but it should give an idea and one conclusion might be that daily maps make sense.

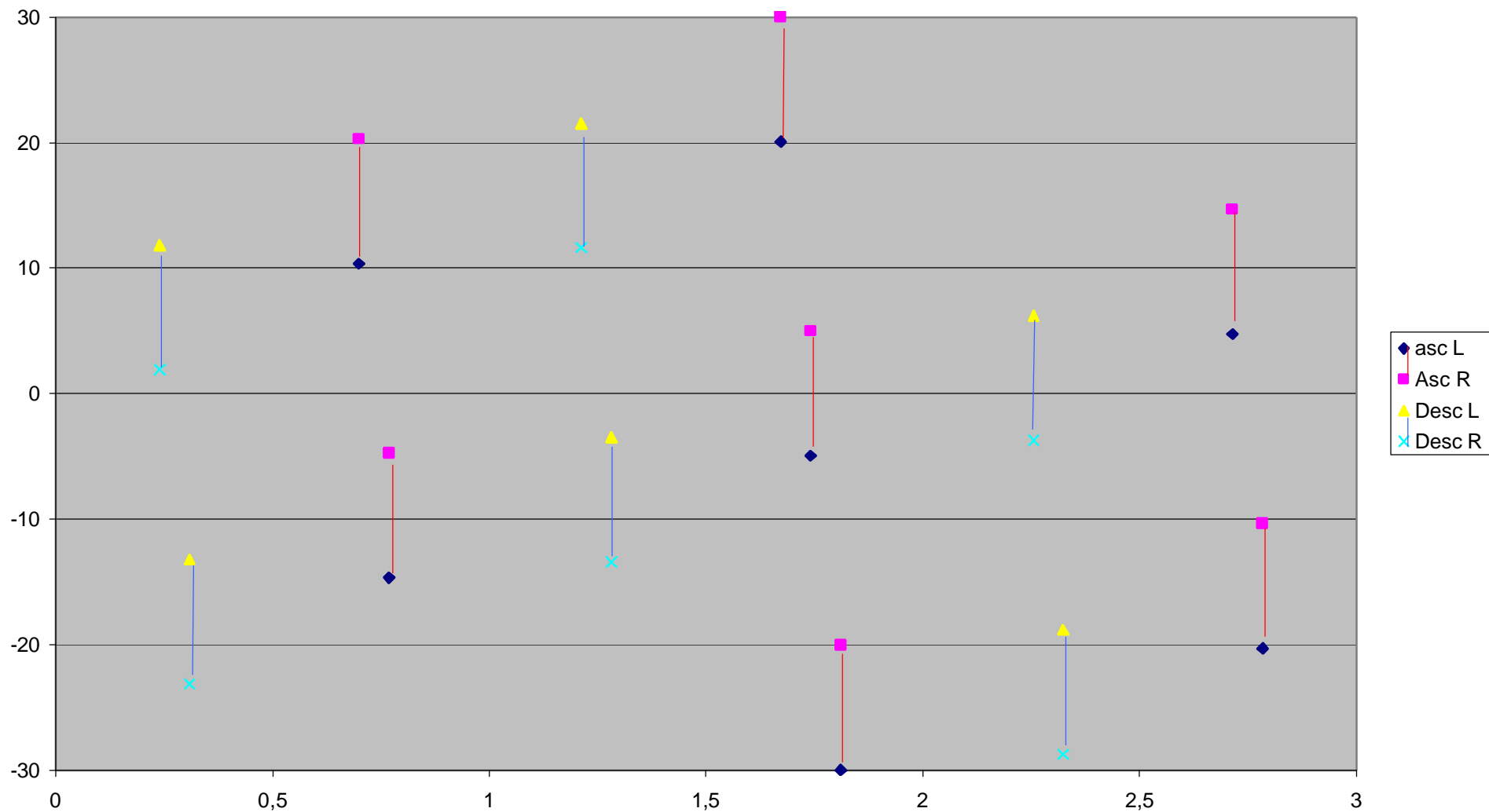
0° N



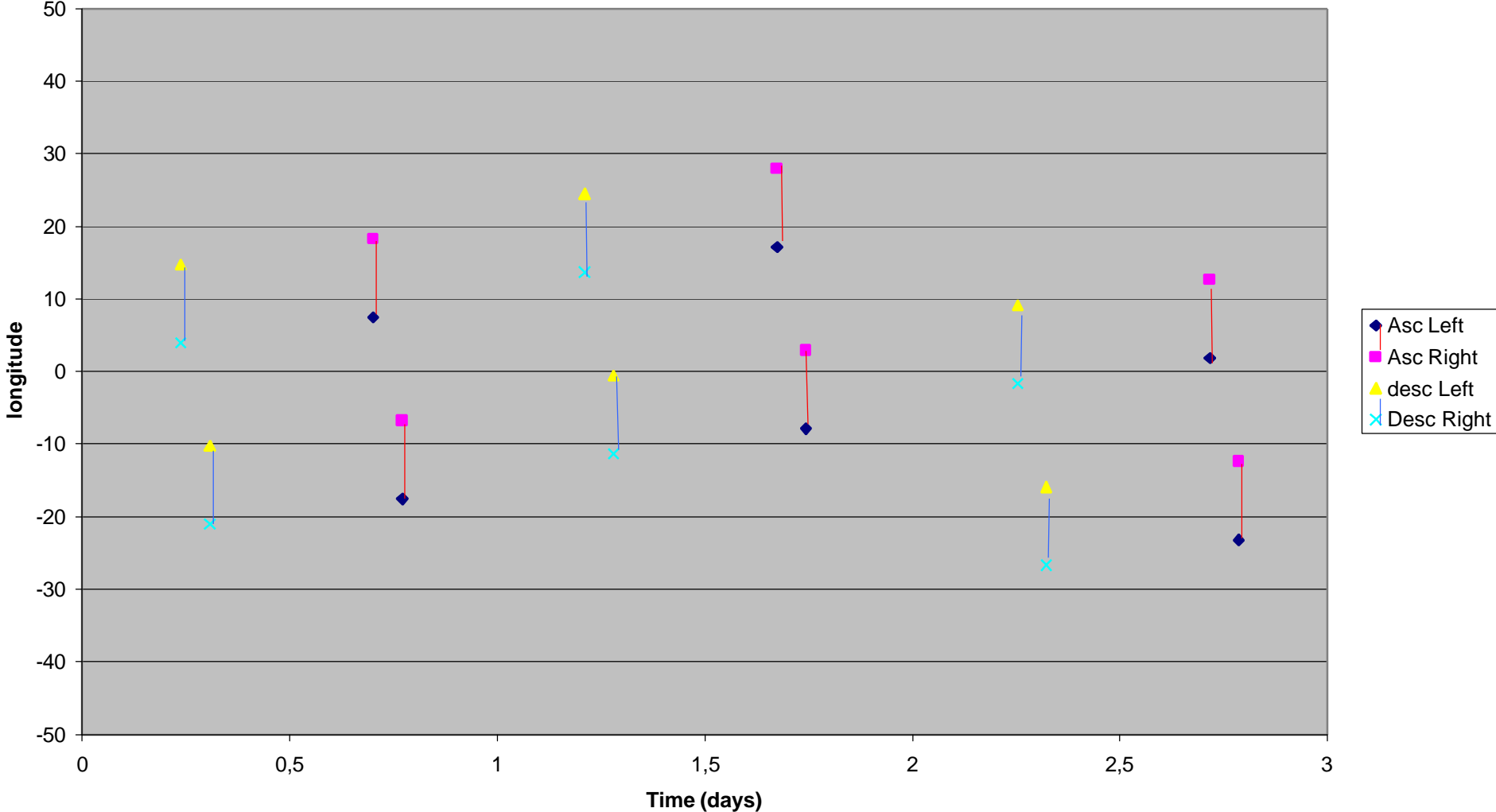
10°



### 20 Deg

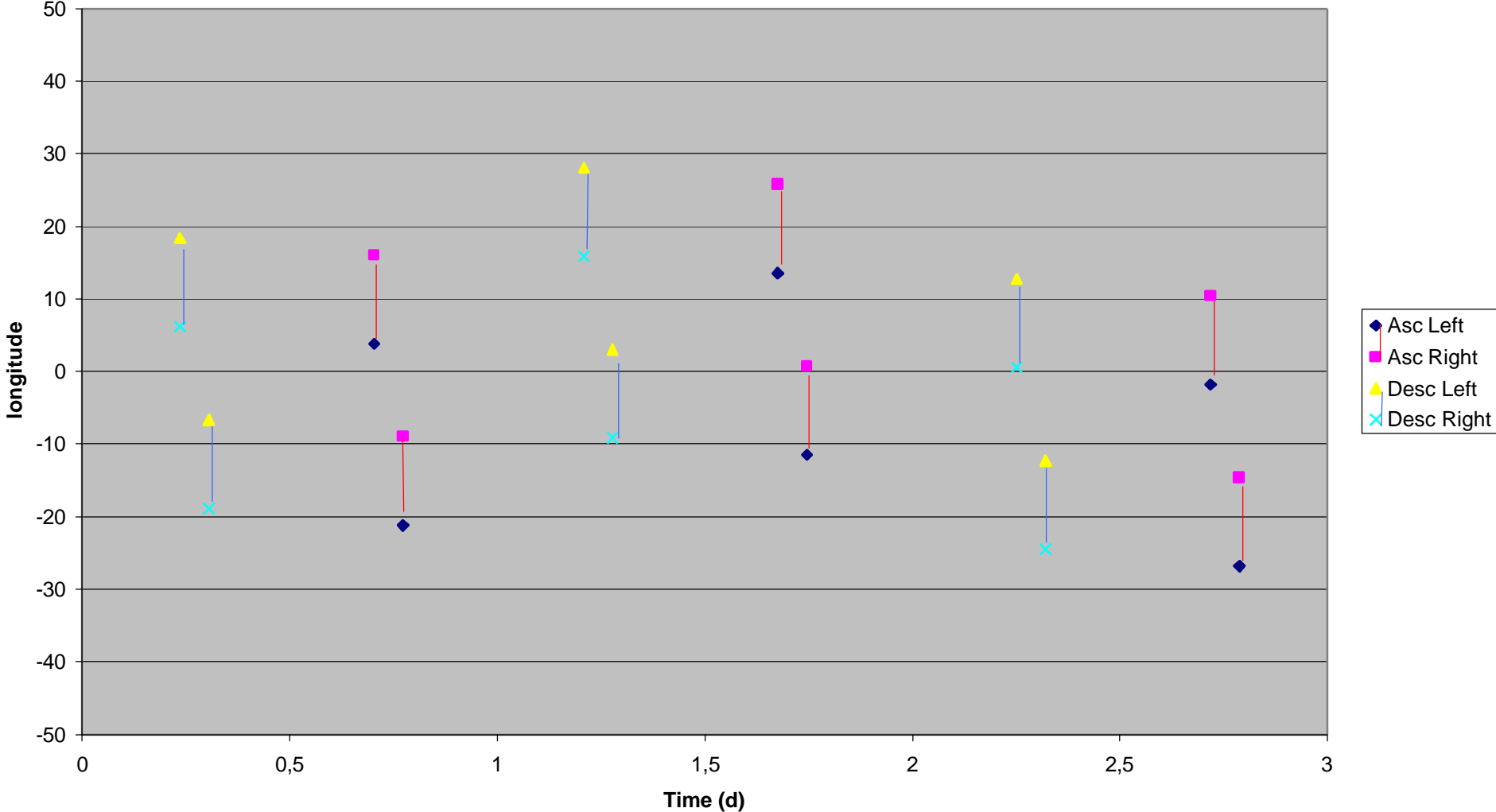


30 deg

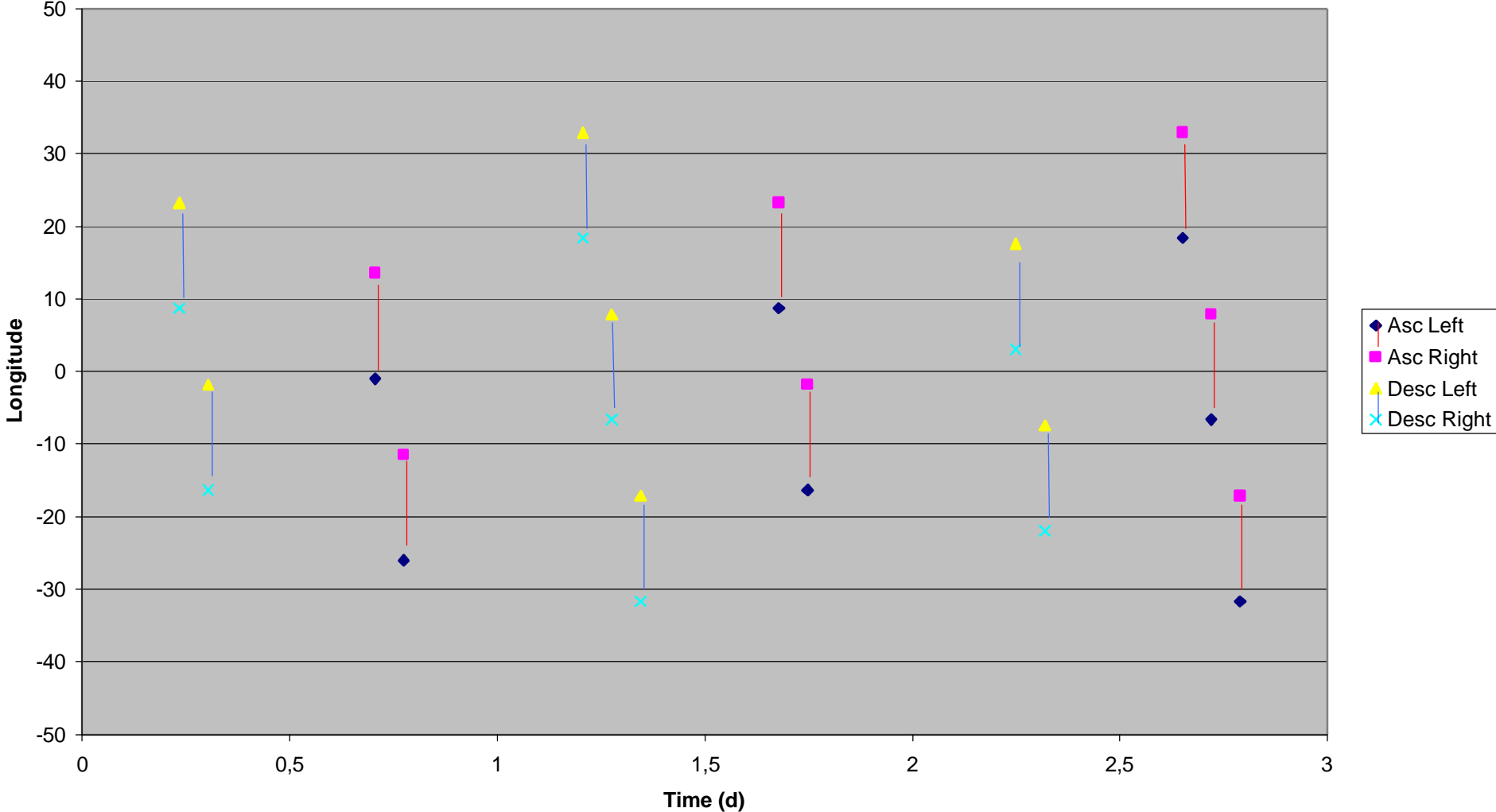




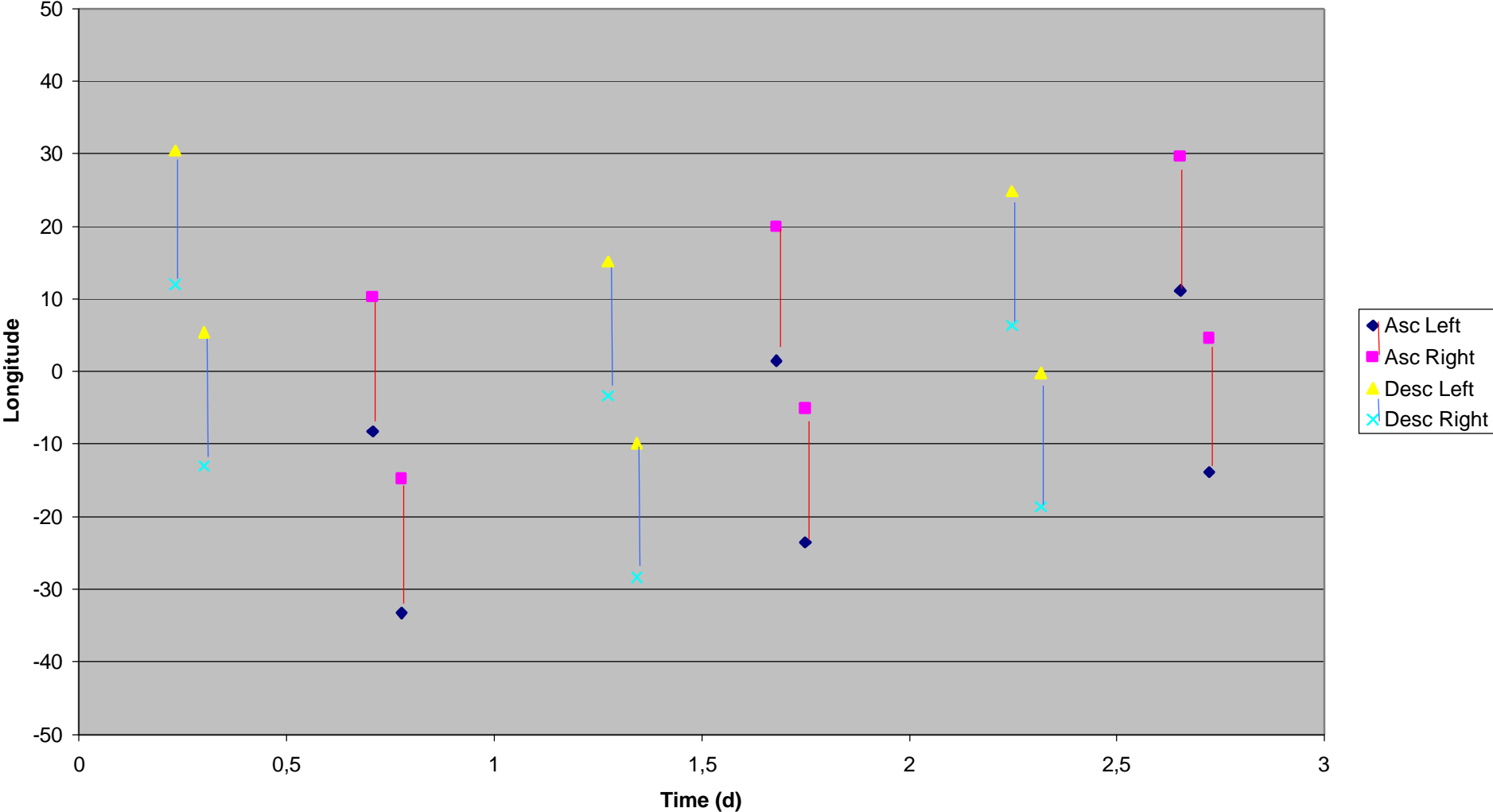
40 deg



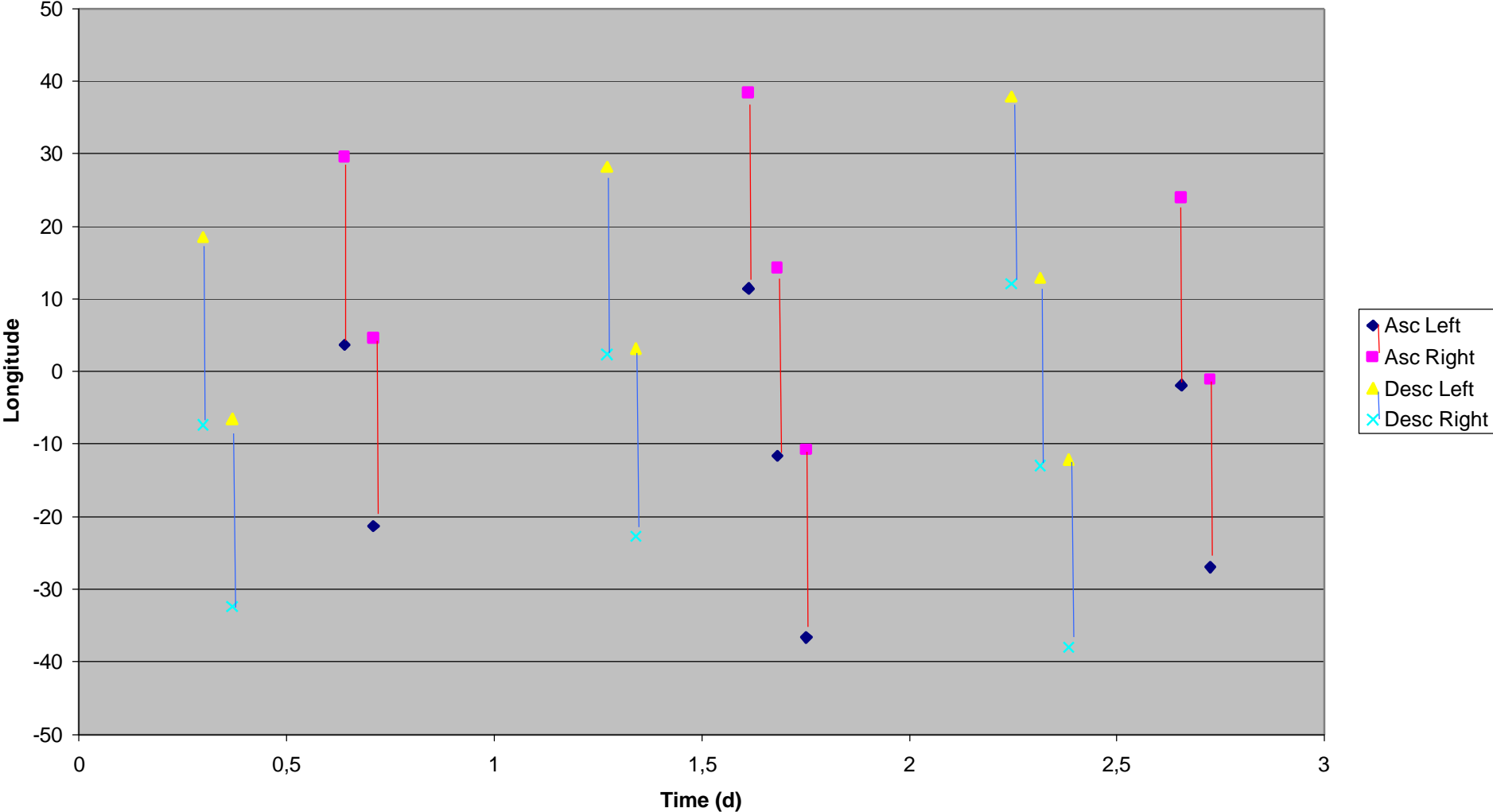
50 deg



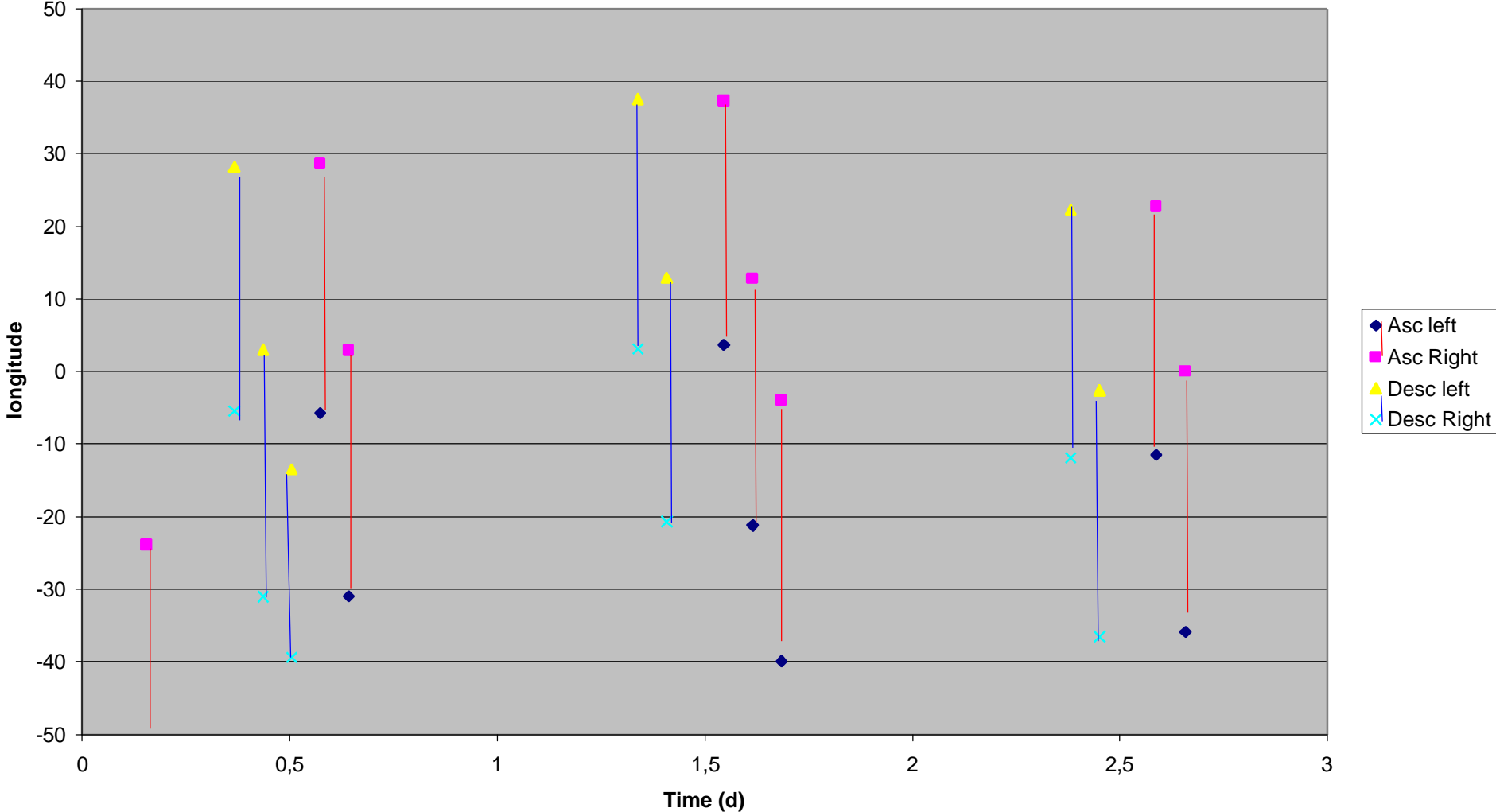
60°



70 deg

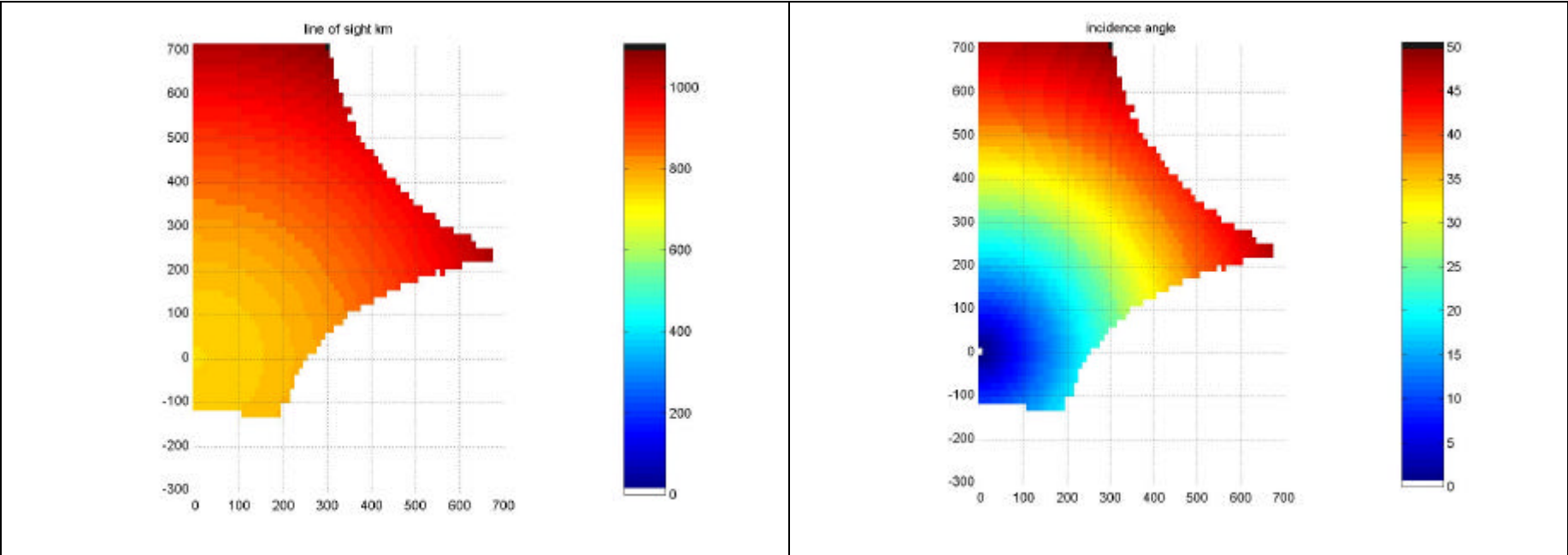


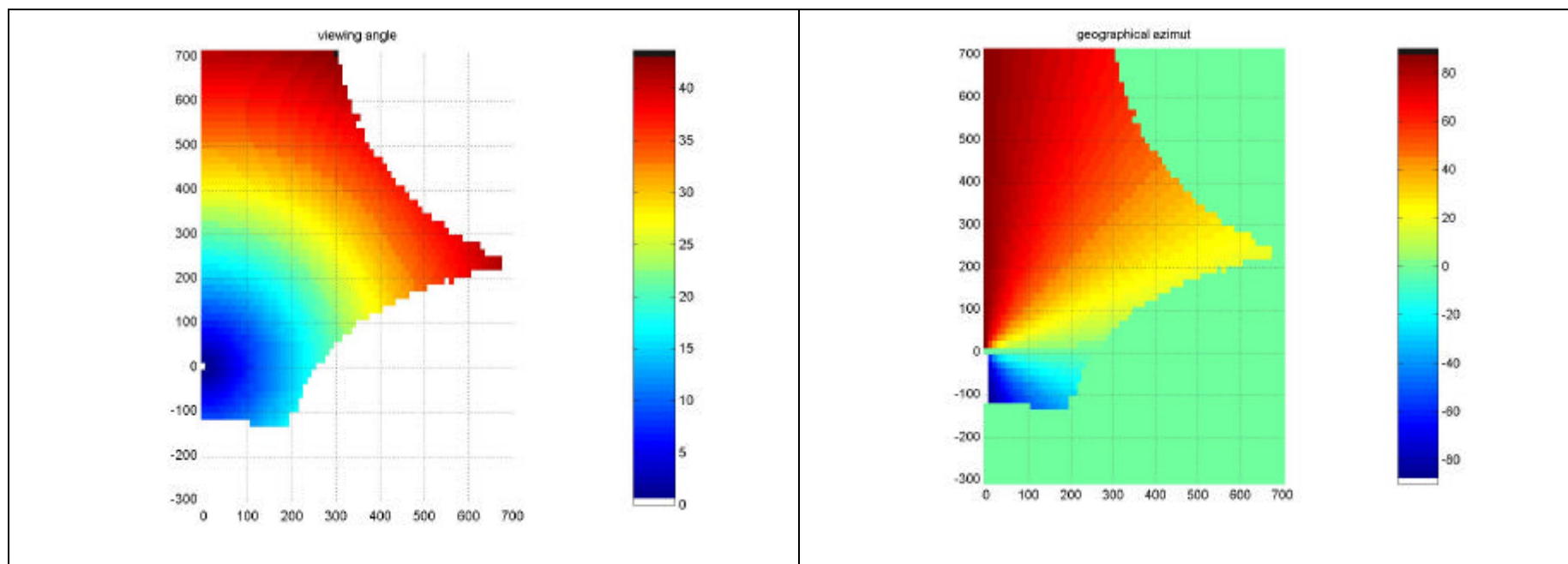
80 deg



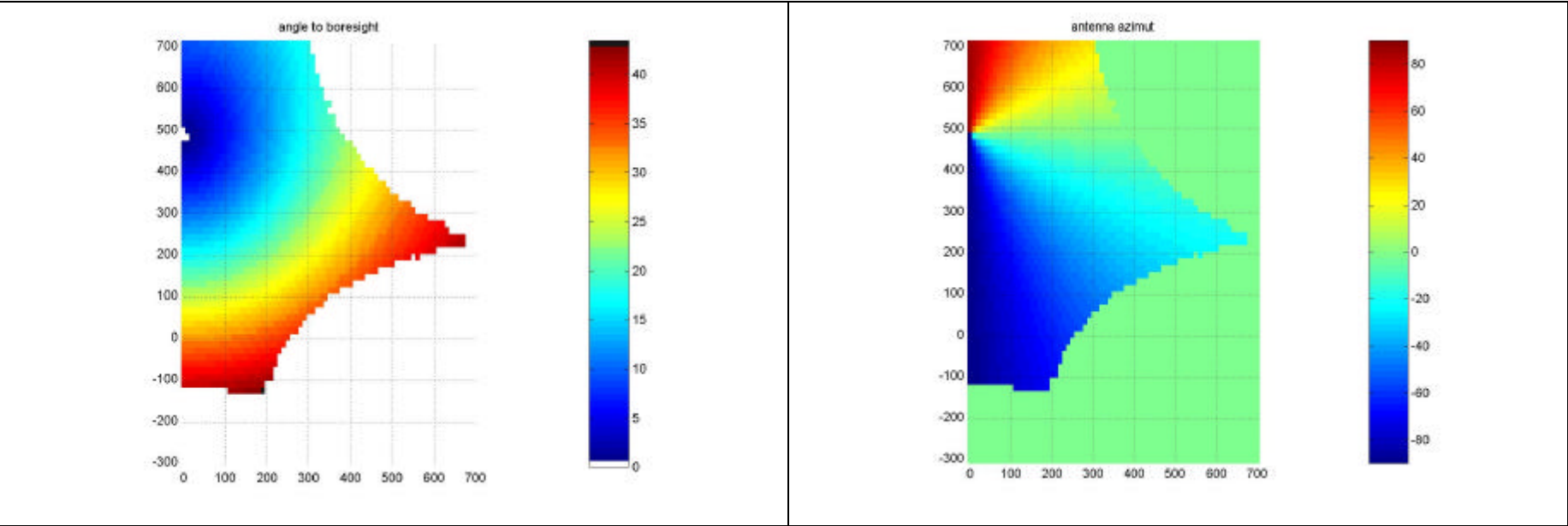
**graphes pour configuration de référence ; nopt71**

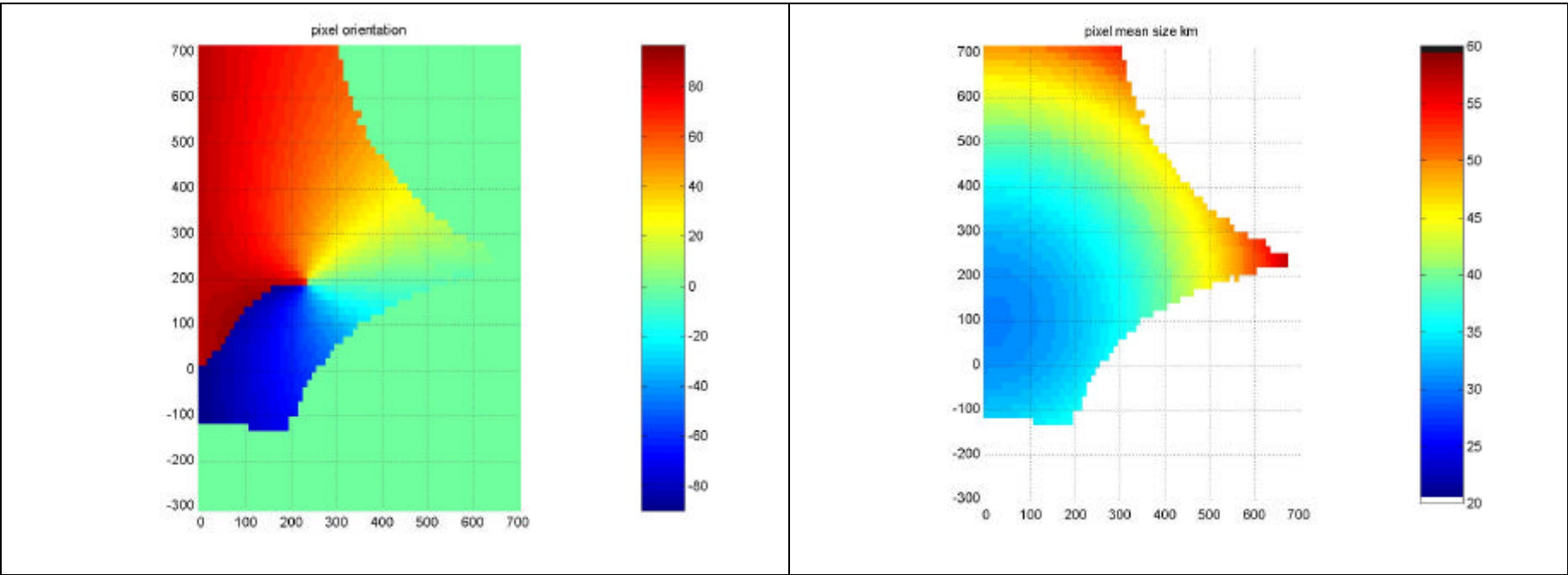
**XX NOPT71** : D=2002 1 9 T=14 59 55 chi=0 dia=0,2 nly,lly,min=99 1 1 jst0=1 0 jan0=0 0 jva0=0 0 jfr0=0 0  
**OPTIONS1** : dx,dy=10 16 dc=0.0077 Rter=6378 vsat=6.6667 (e=.89: 3dBW=32.5, Glin= 8.70)  
orb=14.3; D+D; DX, Dt=200, 10  
**OPTIONS2** : mdB=0.0 sky,sys=3.5 180 com=2 bw=19.0 cbit,frd=1.81 0.820 fwi/ma=0.720 0.80 alf/w=1.00 0.452  
**OPTIONS3** : jdot=1 jnoi=0  
  
**UPDATES** \_\_\_\_: GRID 0= 0.07 PIX 0= 0.01 MSK N= 8.01 D=10.92 T= 0.52 DIAG DIR N= 0.71 INV=DIR t= 0.01  
**CONFIG.** \_\_\_\_: H=755km e,n,L=0.880 21 4.02m tilt,st=32.0 30° Dtet=32.9 Ym, Md=-304 704 16km dt=2.4s sc=1e+000  
**BORESIGHT** : y,z=483 773 SP= 912 arc=484 hor=63.4°(ce=1) - TB DIR N= 1.84  
  
**LOOP sws2** : (1-1); n=1; sws2,4=40 40 dTBax=11510+ 64xT bTB= 0+ 0xT jsta= 1 Tscene=243 243 14 14  
**TARGET** : LAND Ts=293(2.0, bias=0.0) SS,W=35 8(2.00) Ws,Wv=0.10 3.5(0.00) bias2=0.00  
**RESO LOOPS** : jres=1, resm=100; jlrx=1,lrx=3(meanax): resM=1.0 x100; jrem=1,elongM=2.00 - NX LOOP

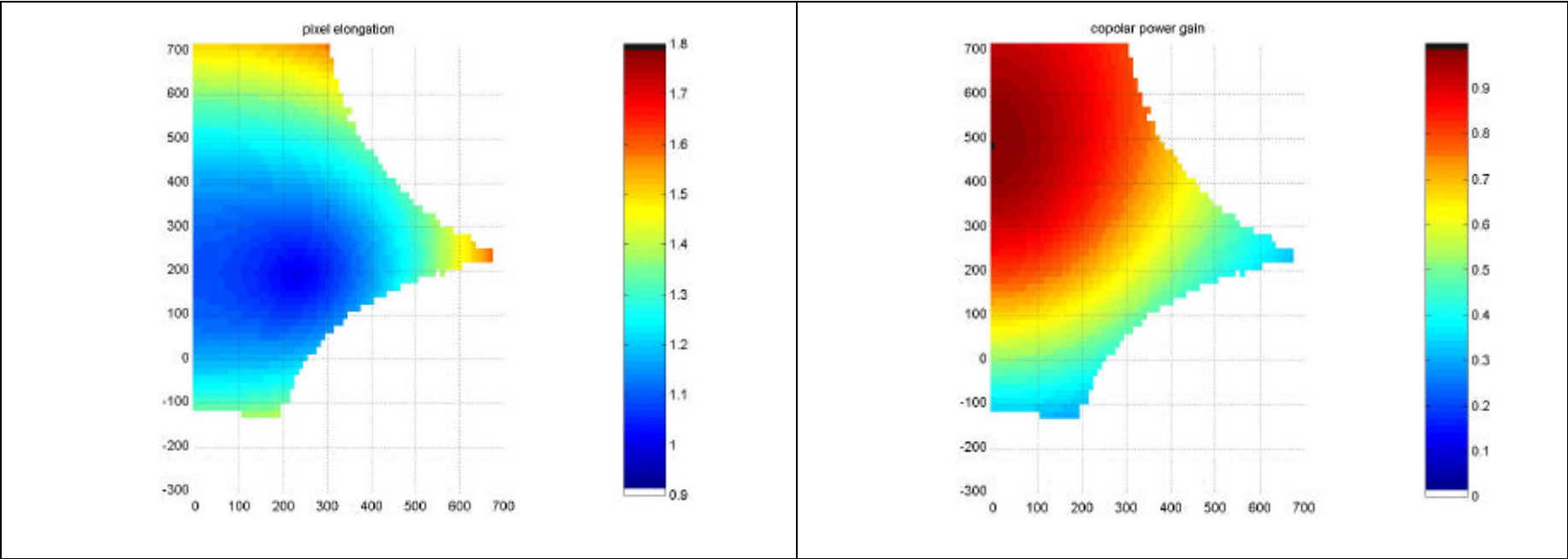


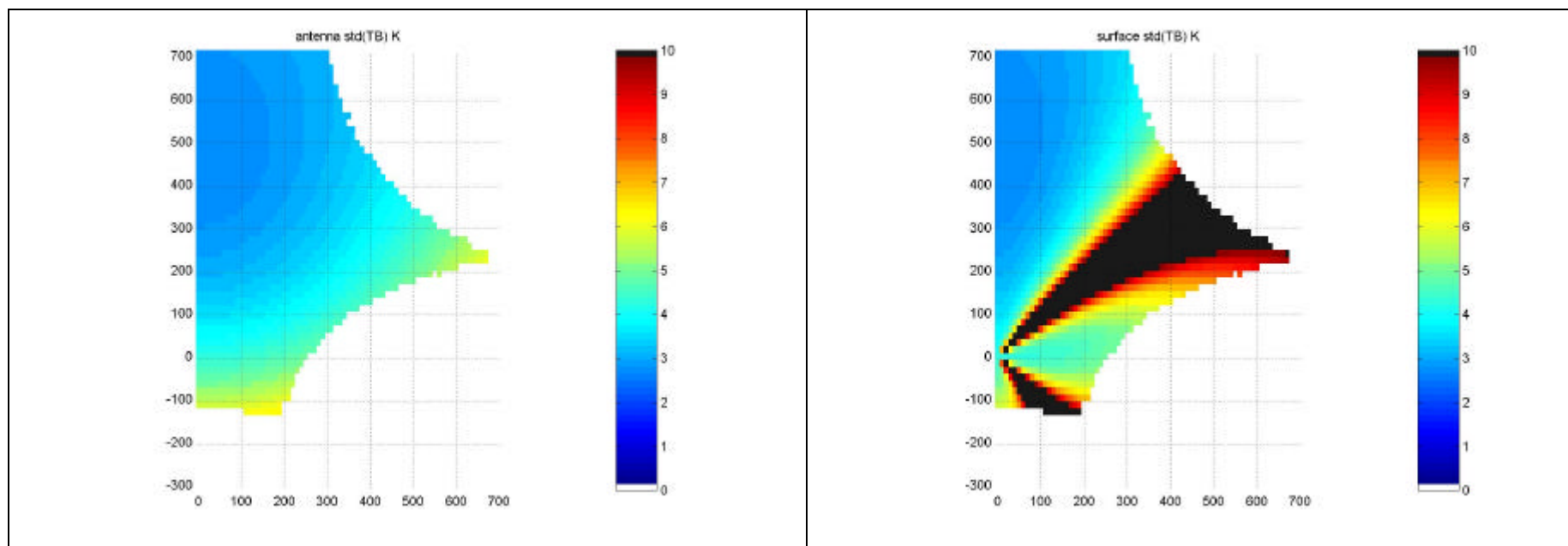












Page: 4

[YHK1] clarify

Page: 4

[YHK2]not clear

Page: 8

[YHK3]dewell line

Page: 8

[YHK4]define

Page: 9

[YHK5]not clear

Page: 10

[YHK6]not clear (science)